The X-ray Crystal Structure of Human β-Hexosaminidase B Provides New Insights into Sandhoff Disease

Timm Maier¹, Norbert Strater¹, Christina G. Schuette²
Ralf Klingenstein², Konrad Sandhoff² and Wolfram Saenger¹*

¹Institut für Chemie Kristallographie
Freie Universität Berlin
Takustr.6, 14195 Berlin Germany

²Kekule-Institut für Organische Chemie und Biochemie der Universität Bonn
Gerhard-Domagk Str.1, 53121 Bonn, Germany

Human lysosomal β-hexosaminidases are dimeric enzymes composed of α and β-chains, encoded by the genes HEXA and HEXB. They occur in three isoforms, the homodimeric hexosaminidases B (ββ) and S (αα), and the heterodimeric hexosaminidase A (αβ), where dimerization is required for catalytic activity. Allelic variations in the HEXA and HEXB genes cause the fatal inborn errors of metabolism Tay–Sachs disease and Sandhoff disease, respectively. Here, we present the crystal structure of a complex of human β-hexosaminidase B with a transition state analogue inhibitor at 2.3 Å resolution (pdb 1o7a). On the basis of this structure and previous studies on related enzymes, a retaining double-displacement mechanism for glycosyl hydrolysis by β-hexosaminidase B is proposed. In the dimer structure, which is derived from an analysis of crystal packing, most of the mutations causing late-onset Sandhoff disease reside near the dimer interface and are proposed to interfere with correct dimer formation. The structure reported here is a valid template also for the dimeric structures of β-hexosaminidase A and S.

Keywords: hexosaminidase; Sandhoff disease; X-ray crystal structure; sphingolipid metabolism; HEXB

Introduction

The human β-hexosaminidases (EC 3.2.1.52) are lysosomal enzymes that degrade the carbohydrate moieties of glycoproteins, glycolipids and proteoglycans through the removal of terminal β-glycosidically linked N-acetylglucosamine or N-acetylgalactosamine residues. They are indispensable in human metabolism for the degradation of gangliosides, an essential class of outer-layer membrane lipids. The dimeric enzymes are composed of two subunits, α and β, that are approximately 60% identical in their amino acid sequence.¹,² The subunits encoded by the genes HEXA and HEXB are synthesized as precursor proteins; processing and subunit assembly in the endoplasmatic reticulum yields three isoforms; β-hexosaminidase A (αβ, HexA), β-hexosaminidase B (ββ, HexB) and β-hexosaminidase S (αα, HexS). Dimerization of two chains is a prerequisite for the formation of catalytically active enzyme. Interestingly, the assembly of HexB from two β subunits is faster than the dimerization of α and β subunits to HexA. Thus, a pool of newly synthesized, unassembled α subunits is retained prelysosomally.³,⁴ Properly folded and assembled enzymes are transferred rapidly to the Golgi apparatus for attachment of the mannose-6-phosphate recognition marker and targeted subsequently to the lysosomes, where final processing produces the mature enzymes.⁵

The wide substrate specificity of the A, B and S-isoforms of β-hexosaminidase is, at least in part, due to a unique active site carried by each subunit: the β-subunit predominantly hydrolyzes neutral substrates, whereas the α-subunit can cleave...
negatively charged substrates. The hydrolysis of ganglioside GM2 by HexA and of the sulfated glycosphingolipid SM2 by HexS requires coactivation by the GM2 activator protein (GM2AP). It lifts the gangliosides out of the membrane and presents their terminal sugar moieties to the water-soluble enzyme.

The particular importance of human β-hexosaminidases is demonstrated by severe inborn errors of metabolism: mutations in the HEXA gene lead to a lack of the isoforms HexA and HexS, cause Tay–Sachs disease; mutations in the HEXB gene lead to a deficiency of isoforms HexA and HexB and cause Sandhoff disease (SD). Both genetic disorders are characterized by a massive accumulation of ganglioside GM2 in neuronal lysosomes, leading to severe and, in most cases, fatal neurodegeneration.

The only structural information on human β-hexosaminidases available so far is a partial homology model based on the X-ray crystal structure of a related bacterial enzyme, Serratia marcescens chitobiase (SmChb), which belongs to family 20 of glycosyl hydrolases. SmChb and other bacterial homologues differ significantly in structure and function from human β-hexosaminidases, being monomeric enzymes of a different substrate specificity as they degrade only water-soluble substrates. Also, the bacterial enzymes do not form ternary complexes with coactivator proteins and substrates, as observed for HexA and HexS. Thus, only limited insights into the unique and functionally decisive features of human β-hexosaminidases can be gained from analyzing the structures of bacterial homologues. To provide a structural basis for the understanding of folding, processing, dimerization and substrate specificity, we have determined the X-ray crystal structure of human β-hexosaminidase B (hHexB) at 2.3 Å resolution in complex with the transition state analogue inhibitor 2-acetamido-2-deoxy-D-glucono-1,5-lactone (β-lactone).

Results and Discussion

Overall structural features

The crystal structure of hHexB, recombinantly expressed in insect cells, was determined by the
MIRAS method and refined to a crystallographic \( R/R_{\text{free}} \)-factor of 19.6%/23.6% at 2.3 Å resolution (Table 1). hHexB folds into two domains, an N-terminal \( \alpha/\beta \)-domain and a central \((\beta,\alpha)_8\)-barrel domain. \( \alpha \)-Helices and \( \beta \)-strands are termed A to R and a to r, respectively; strands in the central \((\beta,\alpha)_8\)-barrel are additionally numbered 1 to 8 in N to C-terminal order (Figure 1(a)). The N-terminal domain consists of a small two-stranded (b,g) and a larger six-stranded \( \beta \)-sheet (c,d,e,f,g,h,a) that is flanked by one short (A) and two long (B,C) \( \alpha \)-helices, the latter forming the largest part of the interface to the central domain, while the opposite side of the large \( \beta \)-sheet is solvent-exposed (Figure 1(b)). The central domain is dominated by the \((\beta,\alpha)_8\)-barrel that carries insertions (blue in Figure 1(a)) near strands 2 (j), 3 (k), 4 (n) and 8 (r) as well as a C-terminal helical extension (R). After strand 2 (j) as well as before and after strand 8 (r), single helices (E, O, P, respectively) are inserted, while the insertion after strand 3 (k) comprises one helix (G) and a two-stranded antiparallel \( \beta \)-sheet (l, m). Two adjacent helices (I, J) are inserted at the C terminus of strand 4 (n). All active-site residues are located in loops on the C-terminal side of the central \((\beta,\alpha)_8\)-barrel (Figure 1(a) and (b)). According to the observed electron density and mass spectrometry, the enzyme is N-glycosylated at Asn84, Asn190 and Asn327 (indicated by linked hexagons in Figure 1(a)), but not at Asn142, while the enzyme purified from human placenta is glycosylated at all four sites.10

**Figure 1.** Fold and homologues of human \( \beta \)-hexosaminidase B. (a) Schematic topogram, \( \alpha \)-helices are represented by circles, \( \beta \)-strands by triangles, numbers indicate the position in protein sequence. Glycosylation sites (hexagons), disulfide bonds (red lines), proteolytic cleavage sites of the mature human enzyme (\( \square \)) and active-site residues (yellow circles) are indicated. (b) Ribbon diagram (stereo), N and C termini and proteolytic cleavage sites of processed HexB (107–122, 311–315) are indicated. In (a) and (b) the N-terminal domain is shown in orange (sheets) and red (helices), the C-terminal domain in green and cyan (strands and helices of central \((\beta,\alpha)_8\)-barrel) and blue (extensions). In (b) the bound \( \delta \)-lactone inhibitor is shown in red. (c) Least-squares superposition of hHexB (blue) with SpHex (green; pdb 1hp5) shown in cartoon representation. The view is identical with that in (b). (d) Least-squares superposition (stereo) of the active-site residues of hHexB (blue) in complex with \( \delta \)-lactone, SpHex (green; pdb 1hp5) with the bound substrate chitobiose and SmChb (gold; pdb 1qbb) in complex with NAG-thiazoline. Residue labels relate to hHexB.

**Structural homology to bacterial enzymes**

The protein fold is highly conserved between hHexB and the bacterial members of family 20
glycosyl hydrolases with known structures; namely, a chitobiase from *Serratia marcescens* (SmChb)\(^9\) and an N-acetylhexosaminidase from *Streptomyces plicatus* (SpHex).\(^1\) Though sequence identity with SmChb and SpHex is only 22% and 27%, respectively, the positions of secondary structure elements of the C-terminal (β,α)\(_8\)-barrel and, unexpectedly, the N-terminal domain are well conserved in hHexB (Figure 1(c)). While SmChb contains two extra domains, extending the protein N and C-terminally, SpHex contains only the domains shared with hHexB. Structural alignments of SmChb/SpHex with hHexB identify 344/349 (71%/72% of residues in hHexB) matching residues with an r.m.s.d. between the C\(_\alpha\) atoms of 1.4 Å/1.5 Å. However, both bacterial enzymes are monomeric and have different substrate specificities compared to the dimeric human β-hexo-

- Catalytic mechanism: the transition state

  The crystal structure of hHex B is in agreement with the suggested conservation of a common
catalytic mechanism with the bacterial hexosaminidases. The family 20 glycosyl hydrolases share a catalytic mechanism with the bacterial hexosaminidase B. The sketch in the upper panel illustrates the two-step double-displacement mechanism: in the first step (1a, 1b) (R = carbohydrate) a substrate-mediated nucleophilic attack on C1 is assisted by the general acid–base catalyst Glu355 and releases the residual carbohydrate R via the formation of a cyclic oxazolinium intermediate. In the second step (2a, 2b) (R = H) a water molecule activated by Glu355 carries out a nucleophilic attack on C1 to open the oxazolinium ring, thereby retaining the initial conformation at C1. In the lower panel, all steps of the mechanism are illustrated by X-ray crystal structures: (a) SmChb together with the substrate chitobiose (pdb 1qbb); (b) hHexB in complex with δ-lactone inhibitor (this work); (c) SpHex in complex with NAG-thiazoline; here, the water position is occupied by a glycerol molecule (pdb 1hp5).

Figure 3. Catalytic mechanism of human β-hexosaminidase B. The sketch in the upper panel illustrates the two-step double-displacement mechanism: in the first step (1a, 1b) (R = carbohydrate) a substrate-mediated nucleophilic attack on C1 is assisted by the general acid–base catalyst Glu355 and releases the residual carbohydrate R via the formation of a cyclic oxazolinium intermediate. In the second step (2a, 2b) (R = H) a water molecule activated by Glu355 carries out a nucleophilic attack on C1 to open the oxazolinium ring, thereby retaining the initial conformation at C1. In the lower panel, all steps of the mechanism are illustrated by X-ray crystal structures: (a) SmChb together with the substrate chitobiose (pdb 1qbb); (b) hHexB in complex with δ-lactone inhibitor (this work); (c) SpHex in complex with NAG-thiazoline; here, the water position is occupied by a glycerol molecule (pdb 1hp5).

Water molecule I (Figure 2(c), I) has an average B value of 33.0 Å² and is fixed in its position by short hydrogen bonds to O² of Asp452 and O¹ of Glu491. Remarkably, this water molecule is coordinated also by the hydroxyl O⁰ of Tyr456 of the other subunit of the dimer (see below) at a distance of 2.7 Å. The coordination sphere would allow interpretation of this water position as a sodium ion. The strong coordination suggests a structural role for this water molecule, similar to that of a second conserved water molecule coordinated by Asp240 and Asp354, which is involved directly in ligand binding by forming a hydrogen bond to O³ of the δ-lactone inhibitor (Figure 2(c), III). Both of these water molecules are buried once a substrate molecule is bound. Water molecule II (Figure 2(c), II) is coordinated and activated by Glu355 and has a higher average B value of 55.3 Å², corresponding to greater mobility or partial occupancy. It is well positioned for an inline attack on the anomeric C1 in the oxazolinium intermediate with an interatomic distance of 3.2 Å to the C1 atom of the transition-state analogue inhibitor δ-lactone. Even with the non-reducing
sugar molecule bound to the −1 subsite, this water molecule is solvent-exposed and water molecules from the bulk solvent could easily enter this position after departure of the residual carbohydrate chain (Figure 2(c)). Therefore, we identify water molecule II as the nucleophilic water molecule attacking the intermediate oxazolinium ring to form the final product in the second step of catalysis (Figure 3).

Crystal packing and quaternary structure

Until now, the dimer structure of human hexosaminidases and its influence on substrate specificity has remained an open question. To identify the physiologically active dimer (for a review, see Gravel et al.5), it is necessary to distinguish between intermolecular crystal packing contacts and the interface of the dimer, which is observed in solution. The crystal packing of hHexB is complex, with the crystal asymmetric unit containing six monomers, organized in three pairs of molecules related by non-crystallographic 2-fold rotation axes. Aside from a few weak contacts involving no more than five residues in each interacting monomer, three intermolecular interfaces are observed, numbered I to III in decreasing interaction strength (Table 2). Interface I involves a larger number of contacting residues (24) than either II (13) or III (11), and an approximately two-fold greater buried surface area per monomer (1368 \( \AA^2 \)) and number of involved hydrophobic residues (eight residues). Although interface I is formed mainly by loop regions, the average \( B \) value of its residues is significantly lower (\( B_{\text{aver}} = 33 \\ AA^2 \)) than for interfaces II (\( B_{\text{aver}} = 40 \AA^2 \)) and III (\( B_{\text{aver}} = 49 \AA^2 \); lower, even, than the average \( B \) value for all protein residues (\( B_{\text{aver}} = 40 \AA^2 \)) (Table 2; Figure 4(a) and (b)). Both the buried surface area and the number of hydrogen bonds between the monomers in interface I are close to the typical values obtained in a survey of 36 permanent protein complexes.14 More amino acid residues involved in formation of interface I are identical (63%) between different mammalian \( \alpha \) and \( \beta \)-subunits than those comprising interfaces II and III (23% and 18%, Table 2). Therefore, only interface I favors the formation of equivalent dimers in HexS \( (\alpha\alpha) \), HexA \( (\alpha\beta) \) and HexB \( (\beta\beta) \). Out of all possible monomer–monomer interfaces, only I positions side-chains of one subunit (e.g. Tyr456, colored blue in Figure 2(c)) in the vicinity of the active site of the other subunit, thus making the regulation of substrate specificity by an interaction partner plausible, as is demonstrated by the different activities of the \( \alpha \)-subunits in HexS and HexA against SM2.15 Thus, the physical strength of the interaction and biological evidence qualifies interface I as a physiologically valid dimer interface, while the others must be regarded as crystal packing interactions.

The dimer interface

The proposed dimer interface is formed mainly by loop regions at the C-terminal face of the \((\beta,\alpha)\)-barrel, as is the active site. The contributing loops are located between strand i and helix D, helices E and F, strand q and helix N, and strand r and helix P. Furthermore, helix P and the C-terminal loop residues from 543 to 550 contribute to the dimer interface. The interface itself is approximately rectangular, with dimensions of 45 \( \AA \) \times 25 \( \AA \). The most striking feature of the interface is a crescent formed by five tyrosine residues (Figure 4(c)). Each of these forms, among other interactions, hydrophobic contacts to one of the five tyrosine residues in the interface region of the other subunit. Contacts of other non-polar side-chains, e.g. those of Leu453, Pro545 and Leu546 (colored dark grey in Figure 4(c)), with one of the five tyrosine residues comprise further hydrophobic interactions. Hydrogen bonding contributes to the dimer interactions to a lesser extent, being observed in only four side-chain–side-chain and one main-chain–side-chain, but no main-chain–main-chain interaction (Figure 4(d)). A considerable number of ordered water molecules is embedded between the two subunits forming hydrogen bonds to residues of both chains.

Anatomy of the substrate-binding site

Though the active center and the binding site of the non-reducing sugar molecule −1 are highly conserved between hHexB and the bacterial homologues, the remainder of the substrate-binding site is significantly different. In the bacterial

---

**Table 2. Properties of intermolecular interfaces I–III observed in hHexB crystals**

<table>
<thead>
<tr>
<th>Properties</th>
<th>I (AB)</th>
<th>II (AC)</th>
<th>III (BC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residues involved in each monomer</td>
<td>211,212, 260</td>
<td>78</td>
<td>71–75</td>
</tr>
<tr>
<td>Residues identical (%c)</td>
<td>452–456</td>
<td>89,92,96,99</td>
<td>123</td>
</tr>
<tr>
<td>Hydrogen bonds</td>
<td>491,492,494</td>
<td>247,249</td>
<td>183,184</td>
</tr>
<tr>
<td>Hydrophobic residues (8 residues)</td>
<td>496,497</td>
<td>500,501,533</td>
<td>262,265,267</td>
</tr>
<tr>
<td>Hydrophobic residues (8 residues)</td>
<td>543–550</td>
<td>269,270,273</td>
<td>552</td>
</tr>
<tr>
<td>No. of residues (9)</td>
<td>24</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>No. of non-bonded contacts (10)</td>
<td>59</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>No. of hydrophobic residues (8 residues)</td>
<td>8</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Hydrogen bonds (10)</td>
<td>10</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>per 100 Å2 area (%)</td>
<td>0.73</td>
<td>1.33</td>
<td>0.63</td>
</tr>
<tr>
<td>Residues identical (%)</td>
<td>63</td>
<td>23</td>
<td>18</td>
</tr>
<tr>
<td>Residues conserved (%)</td>
<td>25</td>
<td>38</td>
<td>36</td>
</tr>
<tr>
<td>Average ( B ) value (Å2)</td>
<td>33 ± 4</td>
<td>40 ± 6</td>
<td>49 ± 10</td>
</tr>
<tr>
<td>Buried surface (Å2)</td>
<td>1368</td>
<td>751</td>
<td>473</td>
</tr>
</tbody>
</table>

---

a. Values are given per interface side.
b. Interface residues contain atoms closer than 3.66 Å to atoms of the interface partner.
c. Hydrophobic residues are Ile, Leu, Val, Met, Tyr, Phe and Trp.
d. Conservation between all human and murine hexosaminidase \( \alpha \) and \( \beta \)-chains.
enzymes, the +1 sugar molecule is bound to a conserved tryptophan residue, Trp685 and Trp408 in the Sm and Sp enzymes, respectively, via hydrophobic stacking interactions. The walls of the sugar-binding crevice are formed by hydrophobic residues in similar positions in SmChB and SpHex; namely, Val493/Val276 (Sm/Sp) on the opposite side and Tyr683/Leu406 on the same side of the +1 sugar ring as compared to the conserved tryptophan residue. No equivalent of these residues binding the +1 sugar ring is present in hHexB (Figure 2(d)). The binding pocket of the human enzyme encloses only the −1 sugar molecule tightly and is wide open around the +1 sugar molecule and following groups in the natural substrates such as glycolipid GA2. A different binding mode for the natural substrates of human hexosaminidases compared with the chitobiose binding of bacterial enzymes is unlikely, as it would interfere with the conserved double-displacement mechanism. Only the binding of the substrate in a distorted sofa conformation, instead
of a relaxed $C_1$ chair one, brings the glycosidic atom O1 in close proximity to the general acid–base catalyst Glu355 (Figure 3). It still remains unclear whether parts of the larger natural substrates beyond the first sugar residue –1 interact with hHexB. In the case of HexA and HexS, it is assumed that most atoms of the substrates, sugar as well as lipid moieties, are in contact with GM2AP15 and only the terminal sugar molecules are bound by hexosaminidases.

Allelic variations of the *HEXB* gene

Genetic defects causing SD include partial gene deletions, smaller insertions and deletions, and individual base substitutions in the *HEXB* gene. Mutations leading to no or highly unstable mRNA are associated with early onset of disease symptoms and a fatal clinical course. Point mutations allowing the production of mRNA and protein with residual activity cause less severe cases of SD, with late onset and slow progression of the disease. Several of these point mutations in hHexB were detected in patients suffering from SD (Table 3).

All residues whose mutations impair protein function are very well ordered in the crystal structure of hHexB, with $B$ values lower than the average for all residues (Figure 4(a)). The functional importance of these residues has been discussed, but with only limited information as the homology model used lacked the N-terminal domain and the C-terminal residues involved in the dimer interface. Moreover, most loops of the dimer interface were modelled incorrectly and protein-protein interactions in the dimer were not known or considered, although human $\beta$-hexosaminidases are active only in the dimeric form. Because most observed mutations are located in regions of dimer contacts or disturb the dimer interface structure indirectly (Figure 4(a)), as will be discussed for the individual mutations in the following paragraphs, the previous model based analysis has proven insufficient.

Of those residues associated with pathogenic mutations, only Pro504 and Arg505 are highly conserved in related eukaryotic and prokaryotic enzymes, while Pro417, Tyr456, Cys309, Cys534 and Ala543 are conserved only in mammalian $\alpha$ and $\beta$-chains. As the active site and the catalytic mechanism of eukaryotic and prokaryotic homologues are virtually identical, this again points to the importance of dimerization properties as a key to understanding the function of hHexB.

### Mutation Ser62Leu: infantile acute SD

The nucleotide exchange C185T resulting in a Ser62Leu mutation in the hexosaminidase $\beta$-chain was found in one allele of an infantile onset SD patient with a partial deletion of the *HEXB* gene in the second allele. No report exists on the expression or activity of the mutant enzyme. Ser62 is located in a sharp $\beta$-turn at the N terminus of the N-terminal $\beta$-strand a. While serine is one of the residues that strongly favors a turn conformation, leucine disfavors it, as reflected by their Chou–Fasman turn parameters of 143 for Ser and 59 for Leu. This suggests an influence of the amino acid substitution Ser62Leu on the overall folding of HexB, as the turn is essential to keep the fold of the N-terminal domain intact.

### Mutation Pro417Leu: subacute or chronic SD due to missplicing

The nucleotide exchange C1214T in exon 11 of the *HEXB* gene leads to the production of an unstable and misspliced mRNA and thus to lowered expression levels of HexB. However, the mutant protein expressed in COS cells also shows only 70% of the specific activity of the wild-type enzyme. This mutation is similar in nature to Ser62Leu, as a proline residue in a $\beta$-turn is replaced by a leucine residue that disfavors loop conformations. The Pro417Leu mutation has been observed in patients together with a Cys309Tyr mutation, which blocks the formation of a disulfide bond, and a common polymorphism Lys121Arg. The combination of these mutations potentially has a stronger effect on the activity of hHexB than a single Pro417Leu mutation.

### Table 3. Point mutations in the hexosaminidase $\beta$-chain detected in Sandhoff disease patients

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Biochemical phenotype</th>
<th>Clinical phenotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>S62L</td>
<td>Unknown</td>
<td>Infantile acute (second allele del50kb)</td>
<td>16</td>
</tr>
<tr>
<td>C309Y</td>
<td>Unknown</td>
<td>Chronic (probably due to second allele P417L)</td>
<td>19</td>
</tr>
<tr>
<td>P417L</td>
<td>70% of normal activity in mutant expressed HexB</td>
<td>Subacute or mild chronic</td>
<td>19,20,43</td>
</tr>
<tr>
<td>Y456S</td>
<td>HexB absent, $\beta$ chains only in HexA, HexA activity: 30% of wild type</td>
<td>Subacute (second allele I207V)</td>
<td>21,44</td>
</tr>
<tr>
<td>P504S</td>
<td>Approx. 20% activity against MUG, mainly in proenzyme form</td>
<td>Chronic (second allele del16kb)</td>
<td>23</td>
</tr>
<tr>
<td>R505Q</td>
<td>Thermolabile</td>
<td>Chronic (second allele null)</td>
<td>22</td>
</tr>
<tr>
<td>C534Y</td>
<td>Unknown</td>
<td>Infantile acute</td>
<td>24</td>
</tr>
<tr>
<td>A543T</td>
<td>Thermolabile</td>
<td>Asymptomatic</td>
<td>26</td>
</tr>
<tr>
<td>K121R</td>
<td>Unaffected</td>
<td>Polymorphism</td>
<td>20</td>
</tr>
<tr>
<td>I207V</td>
<td>Unaffected</td>
<td>Polymorphism</td>
<td>16</td>
</tr>
</tbody>
</table>
Mutation Tyr456Ser: subacute or chronic SD

Two mutations in the hexosaminidase β-chain were observed in a patient suffering from a juvenile onset motor neuron disease, a paternally inherited Ile207Val and a maternally inherited Tyr456Ser mutation. The Ile207Val mutation has been characterized as a common polymorphism not associated with diminished enzyme activity or SD. Tyr456 is located in a loop on the C-terminal side of strand q (strand 7 of the (β,α)8-barrel; Figure 1(a)), close to the active-site residues Tyr450, Asp452 and Leu451, the latter being affected also by the Arg505Gln mutation (see below). Thus, one possible effect of the mutation is a change in loop conformation affecting the active-site geometry, though the distance between Tyr456 and the bound ligand is about 20 Å (Figure 5(a)). As mentioned above, Tyr456 plays an important role in dimerization, where it forms strong hydrophobic interactions with Ile454 and Tyr492 of the other subunit, and is located on the crescent around the dimer interface (Figure 4(c)). These contacts aid in the fixation and positioning of loops on the C-terminal sides of strands q and r (strands 7 and 8 of the (β,α)8-barrel) carrying active-site residues of the other subunit. The importance of Tyr456 for dimerization is confirmed by biochemical data: the absence of active HexB from cultured fibroblasts of an SD patient carrying the Tyr456Ser mutation together with the presence of β-chains in active HexA suggests a defect in homodimerization of the mutant β-chains. Moreover, Tyr456 protrudes as a tyrosine finger towards the substrate bound in the second subunit, such that its hydroxyl group is at a distance of just 4 Å from the closest atom of the inhibitor α-lactone (Figures 2(d) and 5(a)). The Tyr456 hydroxyl Oβ coordinates a structural water molecule, which is coordinated also by Asp452 and Glu491 of the other monomer, as discussed above (Figures 2(c) and 5(a)). Therefore, Tyr456 may indirectly influence the active-site structure of the other subunit and is a potential binding partner for larger and natural substrates. A mutation to the smaller serine would abolish the hydrophobic intermolecular contacts formed by the bulky aromatic tyrosine ring and may affect the stability of the HexB dimer, as well as the geometry of the catalytic site on both subunits of the dimer.

Mutations Pro504Ser and Arg505Gln: heat-labile HexB and chronic SD

Bolhuis et al. identified the mutation Arg505Gln in the β-chain of HexB as the molecular basis for an adult onset form of SD. This variant enzyme was expressed in COS cells and its activity was found to be heat-labile as compared to the wild-type enzyme. Arg505 is located at the N terminus of helix Q and short hydrogen bonds are formed between its guanidyl nitrogen atoms N‡H1 and N‡H2, and the backbone oxygen atom of Leu451. Residue Leu451 is located in a loop at the C terminus of strand q (strand 7 of the (β,α)8-barrel) and its direct neighbors Tyr450 and Asp452 play important roles in substrate binding (Figure 2(a) and (b)). We suggest that the hydrogen bond between Arg505 and Leu451 is essential to stabilize the strongly bent loop conformation around Leu451 and thus to maintain the geometry of the −1 sugar-binding site. The guanidyl N‡H1 of Arg505 forms a short hydrogen bond to the backbone oxygen atom of Arg501 located in helix P. Arg501 itself is involved directly in dimerization: a hydrogen bond is formed between its guanidyl N‡H2 and the side-chain Oe1 of Gln544 in the other subunit (Figures 4(d) and 5(b)). Therefore, Arg505 acts together with Tyr456 as a second direct link between dimerization and active-site structure. A mutation of the neighbouring residue of Arg505, Pro504, to serine results in the expression of heat-labile HexA with reduced activity towards

Figure 5. Overview of residues associated with pathogenic mutations. The hHexB subunit carrying the mutation associated residue(s) is shown in green, the other one in blue. Residues mutated in pathogenic variants are colored grey and labelled red. The cocrystallized δ-lactone inhibitor is shown in red. Only selected residues are shown. (a) Tyr456; (b) Pro504 and Arg505; (c) Cys534 and Ala543.
its natural substrates, but not towards small artificial substrates. This indicates that the mutation affects substrate binding, but not catalytic turnover. 

Pro504 is located at the beginning of a β-turn and adopts a cis-peptide conformation. A serine residue is unlikely to accommodate a similar conformation and will consequently distort the loop conformation. Such alterations would change the orientation of the neighboring C-terminal loop that forms major parts of the dimer interface, as well as the positioning of the adjacent residue, Arg505, discussed above (Figure 5(b)).

**Mutation Cys534Tyr: infantile acute SD**

Kuroki et al. have found a patient with infantile SD to be homozygous for a HEXB mutation resulting in a Cys534Tyr substitution, originally described as Cys522Tyr. Cys534 of the hexosaminidase β-chain forms a disulfide bond to Cys551, which fixes the C-terminal loop to the C-terminal core domain of the protein. Considering one monomer only, the non-disulfide bonded C-terminal loop is expected to be disordered and solvent-exposed as residues 547–552 have no contact to the rest of the protein except the disulfide bond involving Cys551. In the active dimer, these C-terminal residues form an integral part of the dimer interface and all residues from 543 to 550, i.e. to the disulfide bonded Cys551, form interchain contacts. The mutation of Cys534 to tyrosine leaves the C terminus conformationally unrestrained and affects a large portion of the interaction potential between the two subunits of the dimer (Figure 5(c)).

**Mutation Ala543Thr: heat-labile HexB**

Upon incubation at 50 °C for three hours, HexA is inactivated completely, whereas HexB activity remains unchanged. This heat-stability of HexB is commonly used to differentiate it from HexA. However, a heat-labile form of HexB was identified by Navon & Adam, which occurs in different frequencies among various population groups and exists in unaffected persons as well as some SD patients. Narkis et al. identified the underlying mutation G1627A in the HEXB gene as causing an Al543Thr change in HexB. Residue Ala543 is part of the dimer interface and a hydrogen bond is formed between its backbone oxygen atom and O' of Thr496 of the other subunit. Additionally, the neighboring residues in sequence and space, Arg533 and Gln544, both form hydrogen bonds to side-chains of residues of the other dimer subunit, Asp494 for Arg533 and Asn497, Arg501 for Gln544. The change of Ala543 to the bulkier Thr543 is expected to affect the packing of this residue and consequently distort the geometry of the dimer interface, explaining the heat-lability of the resulting HexB variant (Figure 5(c)).

**Conclusion**

We report here the crystal structure of recombinantly expressed, homodimeric human β-hexosaminidase B with a glycosylation pattern nearly identical with that of the enzyme purified from human placenta. On the basis of this crystal structure, the function of individual amino acid residues mutated in patients suffering from Sandhoff disease has been discussed. Most of these residues are involved in formation of the proposed dimer interface directly (Tyr456, Ala543) or indirectly (Pro504, Arg505, Cys534). Residues Tyr456 and Arg505 even act as direct links between the geometry of the active site and the structure of the dimer interface. Remarkably, 20 out of 25 residues in the dimer interface are identical between the β and α-subunits of human β-hexosaminidases and two more residues are conserved, though most of these residues are located in loop regions (data not shown; details of the alignment are available from the authors upon request). The insights concerning dimerization properties are thus not limited to HexB but will also provide a key to the mechanism by which the substrate specificities of HexA and HexS are determined.

The availability of structural details of the complete binding site beyond the −1 sugar-binding site facilitates understanding of the interaction of HexB with larger natural substrates. GM2AP binds to HexA and HexS, and stimulates substrate hydrolysis but does not act on HexB. Based on the crystal structure presented here, residues in the vicinity of the active site on both monomers of the dimer can be identified, which are not conserved between the α and β-chains of β-hexosaminidases. These will be the primary targets of mutational studies to analyze the binding of GM2AP protein to hexosaminidases. Together with the known X-ray crystal structure of GM2AP, models for the ternary complex of natural substrate, GM2AP and β-hexosaminidase can be constructed by computer modelling that will guide future biochemical investigations on the mode of coactivation in glycolipid degradation by β-hexosaminidases. Furthermore, the structure of hHexB will aid in the rational design of hexosaminidase inhibitors recently described as promising drug candidates for the treatment of osteoarthritis. To render such inhibitors nontoxic, they must be as specific as possible. HexS acts on glycosaminoglycans, which are of particular interest in osteoarthritis development. The −1 sugar-binding site of the α-subunits in HexS and HexA are identical, but the different dimerization partners in both enzymes will lead to structural differences in the surrounding of the active sites. These differences can be exploited in the rational design of isoform-specific inhibitors of hexosaminidases. The future potential of structure determination of recombinant hHexB is therefore not limited to an extended molecular understanding of the hexosaminidase related genetic diseases.
Materials and Methods

Crystallization and X-ray data collection

Human β-hexosaminidase B (hHexB) was expressed using a recombinant baculovirus expression system and purified as described.9 The hanging-drop, vapor-diffusion method was used for crystallization. The reservoir solution was 100 mM sodium citrate buffer (pH 5.6), 16% (v/v) ethylene glycol, 10% (v/v) polyethylene glycol 8000. The protein solution contained 10 mg/ml of hHexB in 10 mM sodium citrate buffer (pH 6.0), 100 mM NaCl. The drop was formed by adding 4 μL of protein solution and 1 μL of 5 mM 2-acetamido-2-deoxy-D-glucono-1,5-lactone (α-lactone; Toronto Research Chemicals Inc., Toronto, Canada) to 5 μL of reservoir solution. Crystals of a size of 0.1 mm × 0.1 mm × 0.4 mm were transferred to a cryobuffer (100 mM sodium citrate buffer (pH 5.6), 30% ethylene glycol, 12.5% polyethylene glycol 8000, 0.5 mM δ-lactone) and flash-cooled in liquid nitrogen. For heavy-atom derivatization, crystals were soaked overnight in cryobuffer containing 1 mM K₂PtCl₄ (Pt) or 1 mM HgCl₂ (Hg) prior to flash cooling. Crystals belong to space group P2₁2₁2₁ with cell parameters a = b = 164 Å and c = 245 Å. Crystallographic data were collected at 100 K at BAMline, BESSY (Berlin, Germany) and at ID14EH1, ESRF (Grenoble, France). Details of data collection and reduction are given in Table 1.

Structure solution, refinement and analysis

HKL Suite39 was used for data reduction. Initial phases were obtained with SOLVE30 in a MIRAS protocol using the Pt and Hg datasets in the resolution range 20–4 Å (Table 1). Phases were improved by electron density modification and NCS-averaging with DM31 exploiting the observed 6-fold non-crystallographic symmetry. Phase extension to 2.25 Å resolution with a native dataset yielded well-interpretable experimental maps of extraordinary quality. The model was built with O,32 set yielded well-interpretable experimental maps of the observed 6-fold non-crystallographic symmetry.

Acknowledgements

We thank Claudia Alings for excellent technical assistance in protein crystallization, Judith Weissgerber and Michaela Wendeler for assistance in insect cell culture, Uwe Müller and Steffi Arzt for support at BAMline, BESSY (Berlin) and ID14 EH1, ESRF (Grenoble), respectively. These studies were supported by Deutsche Forschungsgemeinschaft (SFB 284 and Sa196/38-2), a Boehringer Ingelheim Foundation scholarship to C.S., Fonds der chemischen Industrie and Sommerfeld Stiftung.

References


*Edited by R. Huber*

*(Received 4 December 2002; received in revised form 20 February 2003; accepted 21 February 2003)*